Brain Tumor Detection

A Multiresolution approach

Project by :
Hiranmayi Kolambe

November 15, 2024

Introduction

Brain tumors are among the most challenging medical conditions, often requiring early and accurate diagnosis to improve patient outcomes. As medical imaging technology advances, automated methods for tumor detection are becoming critical to assist radiologists in identifying abnormalities in brain scans with greater accuracy and speed. Multi-resolution analysis techniques have shown promise in addressing this need by enhancing image features at different scales, enabling the detection of subtle and complex structures like brain tumors.

Our project focuses on developing an effective multi-resolution analysis approach for brain tumor detection using a combination of Gabor filters, U-Net architecture, and parameter optimization. The primary aim is to capture tumor features across multiple scales and orientations, allowing for a more precise delineation of tumor boundaries and texture characteristics. By leveraging the multi-scale capabilities of Gabor filters and the powerful feature extraction ability of U-Net, we aim to create a robust framework for tumor segmentation.

In the initial phase, Gabor filters are applied to the input brain scan images. Gabor filters are widely recognized for their ability to capture texture details at various scales and orientations, making them well-suited for identifying distinct patterns associated with tumor tissue. By adjusting the filter parameters, we can focus on specific spatial frequencies that highlight tumor regions, thus enhancing the detection process.

Following this, the preprocessed images are fed into a U-Net architecture, a convolutional neural network (CNN) designed for semantic segmentation tasks. U-Net's encoder-decoder structure, with skip connections, enables it to retain high-resolution details while learning hierarchical representations, which is crucial for accurately segmenting tumor boundaries. We train the U-Net on a dataset of annotated brain MRI scans, allowing the network to learn the distinguishing features of tumor and non-tumor regions.

Finally, we conduct a systematic parameter tuning to optimize the performance of both Gabor filters and the U-Net model. This includes adjusting the scale and orientation parameters of the Gabor filters as well as modifying key U-Net parameters such as learning rate, batch size, and the number of filters. Our objective is to achieve a high level of segmentation accuracy while minimizing false positives and false negatives, thereby enhancing the model's reliability for real-world applications.

In summary, this project explores a multi-resolution approach to brain tumor detection by integrating Gabor filtering with a U-Net-based segmentation model. Through parameter optimization, we aim to refine the sensitivity and specificity of the system, ultimately contributing to the advancement of automated brain tumor diagnosis techniques. This approach holds potential not only for brain tumors but also for broader applications in medical image analysis where multi-scale feature detection is essential.

Gabor Filters

Gabor filters are a family of linear, frequency, and orientation-selective filters. They have been chosen for texture analysis applications due to the evidence from psychophysical research which indicated that the human brain performs a frequency analysis of images. The class of Gabor functions was first introduced by Gabor and was extended to two dimensions by Daugman. Daugman showed that Gabor filters are optimal in a way that they minimize the product of effective areas occupied in the 2D space and frequency domains. These filters can be described in terms of a sinusoidal plane wave of some spatial frequency and orientation within a two-dimensional Gaussian envelope.

Gabor filters have been used for feature discrimination using phase signatures (or phase congruency [PC]). Phase-based feature detection has been studied well since the publication of the Local Energy Model of feature detection. The computation of local PC uses a pair of quadrature filter banks, log-Gabor filters, and can be easily extended to N-dimensions using the monogenic signal representation. A series of orientable 2D filters can be constructed by extending a log-Gabor function into 2D. PC spatial localization can be improved when calculated via wavelets. These studies have given a number of phase-based feature detection algorithms. Mulet-Parada and Noble suggested an intensity-amplitude invariant approach using a phase-based feature detection method. When applied to echocardiographic image sequences, the algorithm takes advantage of the temporal inconsistency of speckle to detect the acoustic boundaries.

Gabor filters represent a Gaussian envelope modulated by a sinusoidal plane wave. There filters can be defined in the spatial domain by the following expression (Štruc & Pavešić, 2010) (Shen et al., 2007):

$$\psi(x, y)= \frac{f_u^2}{\pi k \eta} e^{\left((f_u^2/k^2)x'^2+(f_u^2/\eta^2)y'^2\right)} e^{j2\pi f_u x'}$$

Where, $x' = x\cos\theta_v + y\sin\theta_v$

$y' = -x\sin\theta_v + y\cos\theta_v$

$f_u = f_{max} / 2^{(u/2)}$

$\theta_v = v\pi / 8$

f_u and θv refer to the center frequency and orientation, respectively. k and η represent the ratio between the center frequency and the size of the Gaussian envelope. In this paper, the values of the parameters used are k = η = 2^.5 and fmax=0.25, which are the most commonly used as in (Štruc and Pavešić, 2010, Shen et al., 2007). Also, in order to extract multi-scale and multi-orientation features from the face image, in this study, a bank featuring filters composed of five scales (u=0,1,......,4) and eight orientations (v=0,1,......,7) were constructed. Given I(x,y) the grey-scale face image, the process of features extraction can be defined as:

$$G_{u,v}(x, y) = I(x, y) * \psi_{u,v}(x, y)$$

where Gu,v(x,y) are the complex filtering output.

Gabor filters excel in capturing image texture features. In addition to texture features, the Gabor filter can also recognize useful features in the sample image and is sensitive to changes in image brightness, contrast, etc., thus giving the model greater robustness. The Gabor function consists of a Gaussian function and a cosine function, whose expression is in complex form and consists of a real part and an imaginary part, which are orthogonal to each other. Among them, the filtering of the real part can be used to smooth the image, while the filtering of the imaginary part is used for edge detection. Therefore, based on their excellent performance, Gabor filters are widely used in many fields such as texture analysis, image classification, and action recognition.

$$G(x,y,z;\theta,\phi,\sigma,\gamma,\lambda) = G_{re}(x,y,z;\theta,\phi,\sigma,\gamma,\lambda) + jG_{im}(x,y,z;\theta,\phi,\sigma,\gamma,\lambda)$$

$$G_{re}(x,y,z;\theta,\phi,\sigma,\gamma,\lambda) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2 + \gamma^2 z'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda} + \phi\right)$$

$$G_{im}(x,y,z;\theta,\phi,\sigma,\gamma,\lambda) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2 + \gamma^2 z'^2}{2\sigma^2}\right)\sin\left(2\pi\frac{x'}{\lambda} + \phi\right)$$

$$x' = x\cos(\theta) + y\sin(\theta) + z\phi$$

$$y' = -x\sin(\theta) + y\cos(\theta)$$

$$z' = z\cos(\theta) - y\phi$$

$$b = \log_2 \frac{\frac{\sigma}{\lambda}\pi + \sqrt{\frac{ln2}{2}}}{\frac{\sigma}{\lambda}\pi - \sqrt{\frac{ln2}{2}}}$$

$$\frac{\sigma}{\lambda} = \frac{1}{\pi}\sqrt{\frac{ln2}{2}}\frac{2^b + 1}{2^b - 1}$$

Descriptions of all parameters of the filter.

Parameters	Descriptions
θ	Parallel stripe direction of the filter, value range [0~360°]
ϕ	Offset phase of the filter, value range [−180~180°]
σ	The standard deviation of the Gaussian function that determines the size of the filter space. This value is related only to the semi-response spatial frequency bandwidth, when $b = 1,\ \sigma = 0.56\lambda$.
γ	The spatial aspect ratio, used to adjust the shape of the Gabor filter in the x and y directions. When the value is 1, it is circular; when the value is less than 1, the filter shape is elongated with the direction of the parallel stripes.
λ	Spatial frequency-dependent wavelengths associated with sinusoidal components
b	Half-response spatial frequency bandwidth, related to σ/λ, is a positive real number, generally taken as 1.

U-Net architecture

The U-Net architecture is a type of convolutional neural network (CNN) specifically designed for image segmentation tasks. It was first introduced by Olaf Ronneberger et al. in 2015, primarily for biomedical image segmentation. The U-Net model is structured to capture both local and global features, enabling precise segmentation of images into regions of interest. It primarily relies on convolutional layers. Convolutional layers perform a convolution operation, where a kernel (or filter) slides over the input, applying a mathematical operation to extract features.

For an input image I(x,y)I(x,y) and a kernel K(a,b)K(a,b), the convolution operation C(x,y)C(x,y) can be represented as:

$$C(x,y) = \sum_a \sum_b I(x+a, y+b) \cdot K(a,b)$$

In U-Net, these convolutional layers capture local features, such as edges and textures, in the downsampling path, which are then used to reconstruct the segmented output in the upsampling path. The encoder part of the U-Net progressively reduces the spatial dimensions of the image while increasing the depth (number of feature maps), enabling it to capture abstract features at various scales. The downsampling operation typically consists of two main layers:

Convolutional Layer: Extracts features as described above and Max Pooling Layer: Reduces the spatial dimensions by taking the maximum value in a pooling window (usually 2×22×2), effectively downsampling the image. If the input size is H×WH×W with depth DD, after a max pooling operation with stride 2, the new dimensions become H2×W2×D2H×2W×D. The decoder path is symmetric to the encoder and is used to reconstruct the segmentation map. The upsampling path consists of: Transposed Convolution (Deconvolution): Increases the spatial dimensions of the feature maps by reversing the convolution process. For a convolution kernel of size k×kk×k and stride ss, the output dimensions are increased by a factor of ss and Concatenation: U-Net employs skip connections by concatenating corresponding feature maps from the encoder to the decoder. This helps in retaining fine-grained details lost during downsampling. One of the key elements in U-Net is the skip connection. For each level ii, feature maps from the encoder are concatenated with those in the decoder. Let Fienc represent the encoder feature map at level ii and Fidec be the decoder feature map. The skip connection is represented mathematically as:

$$F_i^{\text{concat}} = \text{Concatenate}(F_i^{\text{enc}}, F_i^{\text{dec}})$$

These skip connections allow U-Net to retain spatial information, enabling precise localization in segmentation. The U-Net is trained using a loss function that quantifies the difference between the predicted segmentation and the ground truth. Common loss functions for U-Net include: Cross-Entropy Loss: For pixel-wise classification, cross-entropy loss is commonly used. For each pixel, it is defined as:

$$L_{\text{CE}} = -\frac{1}{N} \sum_{i=1}^{N} y_i \log(p_i)$$

where yi is the ground truth label and pi is the predicted probability for pixel i, and N is the number of pixels. Dice Loss: Dice loss, based on the Dice coefficient, is widely used in segmentation tasks to handle imbalanced classes:

$$L_{\text{Dice}} = 1 - \frac{2\sum_{i=1}^{N} y_i p_i}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} p_i}$$

This loss function is particularly effective for segmentation as it emphasizes overlap between the predicted and actual regions. Given an input image X∈RH×WX∈RH×W, the U-Net applies several convolutional and pooling operations in the encoder to produce a latent feature representation. After the encoding step, the representation is upsampled in the decoder to output the segmentation map Y∈RH×WY∈RH×W.

Let:

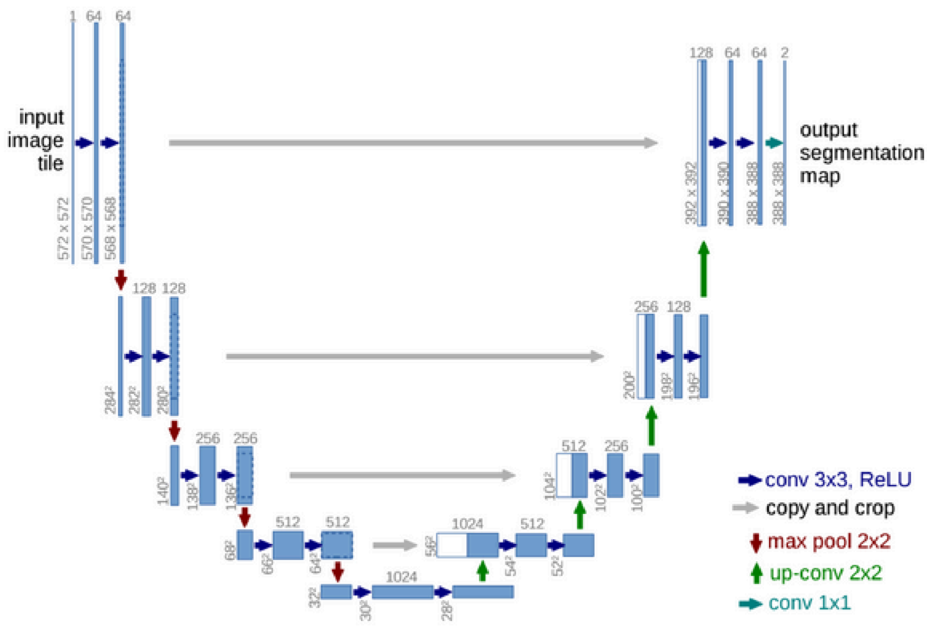
- fenc(·)fenc(·) be the downsampling (encoder) function.
- fdec(·)fdec(·) be the upsampling (decoder) function.

Then, U-Net's forward pass can be represented as:

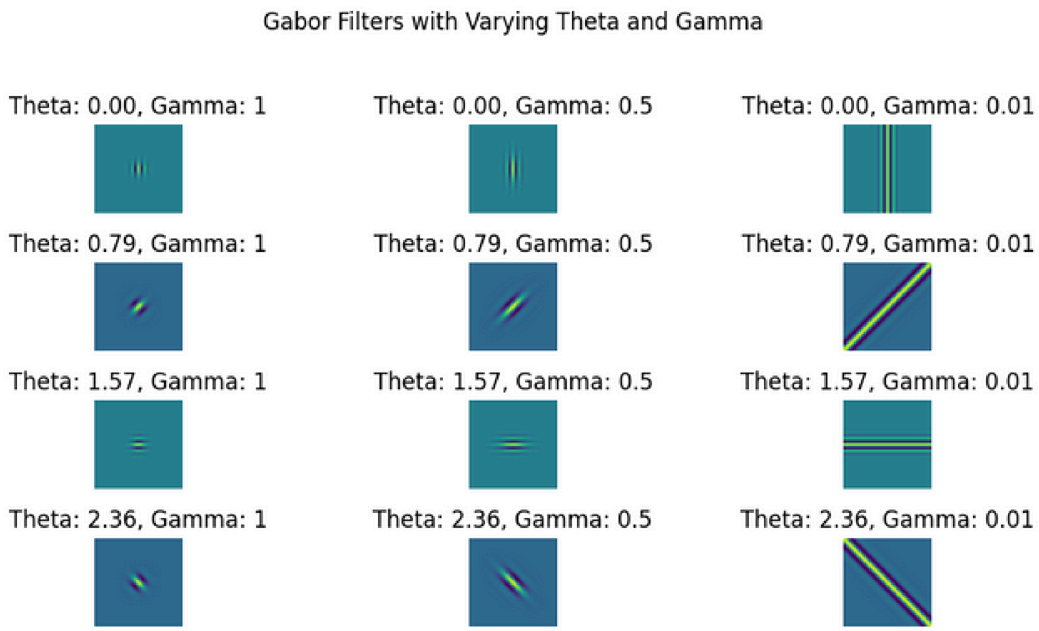
$$Y = f_{\text{dec}}(f_{\text{enc}}(X))$$

where fenc includes both convolution and max-pooling operations, and fdec includes transposed convolutions and concatenations with encoder features via skip connections. This mathematical framework underpins the U-Net's effectiveness in image segmentation, especially in applications requiring high spatial accuracy, like medical imaging.

The advantage of PC is that it is invariant to brightness and contrast, therefore, robust to variations in images, and can be extended to N-dimensions. Some researchers have claimed that the local image phase is more robust than the intensity gradient for acoustic boundary detection. The disadvantage is directly related to the advantage, namely its high sensitivity to noise.



U-netarchitecture(examplefor32x32pixelsinthelowestresolution).Eachblue boxcorrespondstoamulti-channel featuremap.Thenumberof channels isdenoted ontopof thebox.Thex-y-size isprovidedat the lower leftedgeof thebox.White boxesrepresentcopiedfeaturemaps.Thearrowsdenotethedifferentoperations.

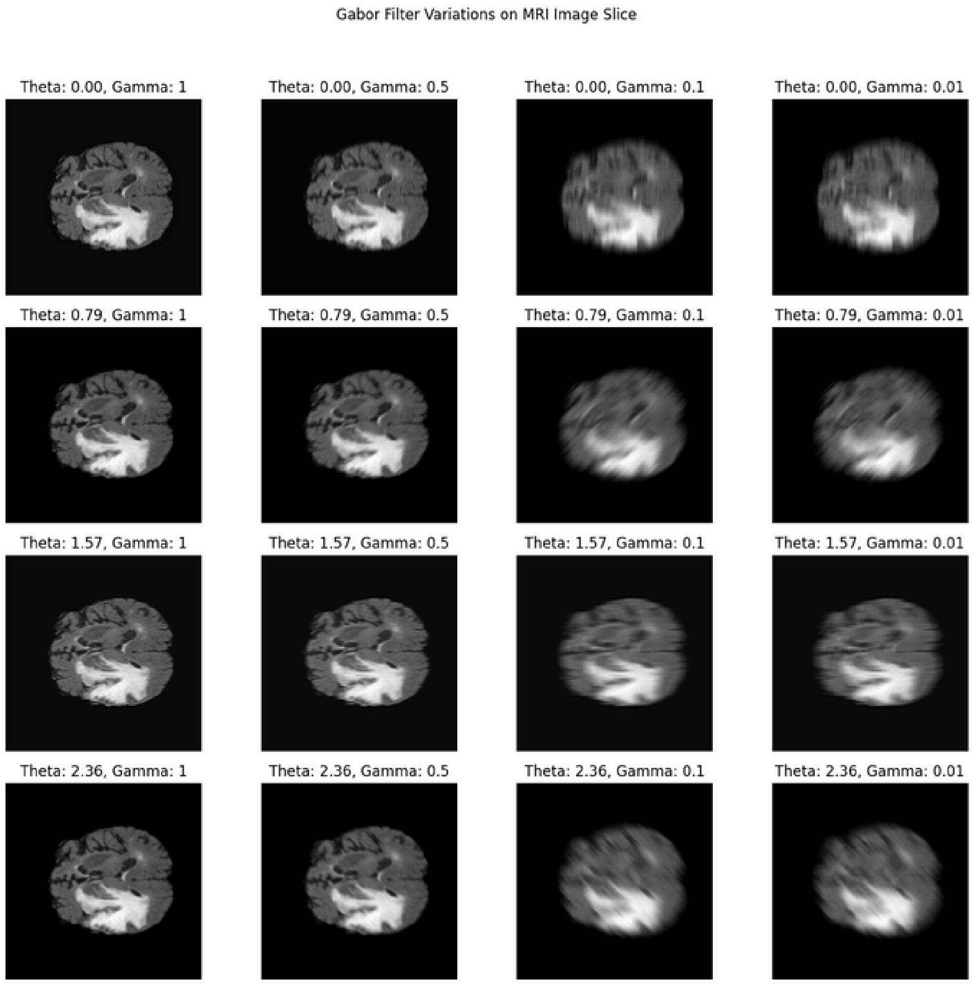


Gabor filters have also been used in ultrasound for edge detection and segmentation in various fields of application, such as kidney imaging, liver sonography, prostate cancer, etc. The above image is a representation of a set of Gabor filters with varying values for the parameters theta (orientation) and gamma (aspect ratio), while keeping other parameters constant: kernel size (ksize) at 15, sigma (the Gaussian envelope's standard deviation) at 0.8, and phi (phase offset) at 0. This exploration of Gabor filters serves to illustrate how different parameter configurations can influence feature detection in image analysis, a useful approach in applications such as texture recognition, edge detection, and medical image segmentation.

The theta parameter, which defines the orientation of the filter, is adjusted in each row of the image. Theta values are approximately set at 0.00, 0.79, 1.57, and 2.36 radians, corresponding to orientations of 0° (horizontal), 45°, 90° (vertical), and 135°, respectively. By altering theta, the Gabor filter becomes sensitive to patterns aligned with these orientations, enabling the detection of directional textures within an image. This orientation sensitivity is particularly beneficial in applications like brain tumor detection, where identifying edges and textures at various angles can assist in highlighting abnormal features in medical images.

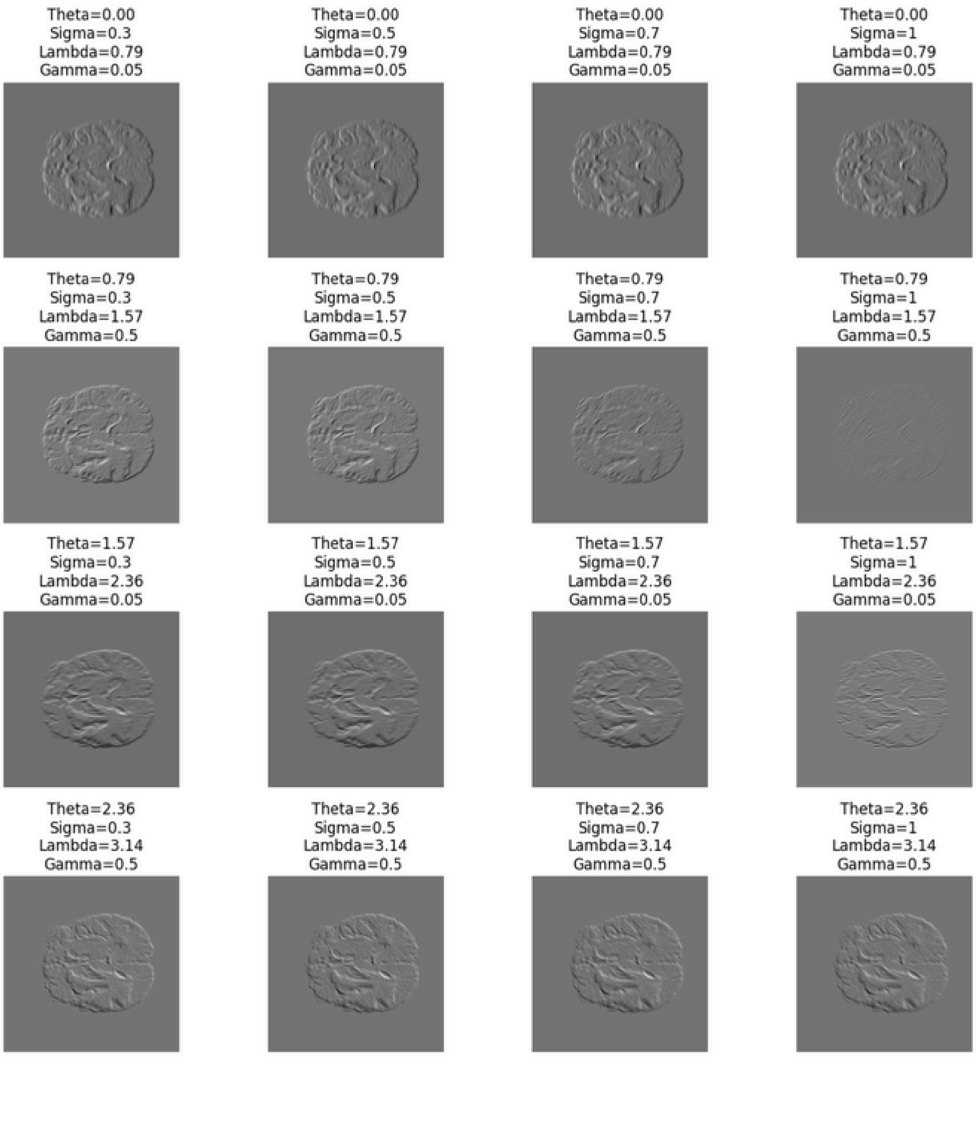
The gamma parameter, representing the aspect ratio, controls the filter's sensitivity to structures of varying thickness. In the image, gamma values are set at 1, 0.5, and 0.01 across the columns. Gamma = 1 produces a more circular Gaussian envelope, making the filter sensitive to broader, more circular features within the chosen orientation. As gamma decreases, the filter becomes elongated, with gamma = 0.5 creating an elliptical shape, making it more responsive to medium-width patterns. At gamma = 0.01, the filter takes on an almost line-like shape, highly sensitive to fine, narrow edges in the specified direction. This degree of control allows for multi-resolution analysis, where filters with different gamma values can capture both coarse and fine details, useful for applications requiring precision in edge and texture detection.

By keeping ksize, sigma, and phi constant, this image isolates the effects of theta and gamma variations, illustrating how these two parameters influence the spatial frequency response and orientation selectivity of the Gabor filters. This selective sensitivity is integral to applications in image processing and computer vision, where identifying features across multiple scales and orientations is often necessary. In medical imaging, for instance, this technique allows for the detection of various tumor textures and boundaries, contributing to more accurate segmentation and diagnosis.



The purpose of these Gabor filter variations is to highlight structures in the MRI slice across multiple scales and orientations, which is particularly valuable for tasks like identifying edges, textures, or regions of interest that may be indicative of abnormalities such as brain tumors. The above matrix of filtered images illustrates how combining different theta and gamma values enables multi-resolution analysis, capturing both coarse and fine details at various orientations. In brain MRI analysis, this approach can help isolate critical regions, such as tumor boundaries or tissue textures, that may require further investigation. By adjusting theta and gamma, Gabor filters provide a versatile tool for enhancing feature visibility in complex medical images, supporting more accurate diagnosis and assessment of pathological conditions.

The filtering performance of a Gabor filter is mainly determined by the size of the convolution kernel. Different convolution kernel sizes may have different effects on the Gabor filter. If the edge length of the convolution kernel is larger than the wavelength, it has no effect on the result of filtering. If the edge length of the convolution kernel is smaller than the wavelength, then the entire waveform is not fully included in the convolution calculation, resulting in poor filtering of the waveform edges. It can be concluded that the edge length of the convolution kernel of the Gabor filter must be greater than the wavelength to ensure that the entire waveform is completely contained in the convolution kernel to maximize the filtering effect of the Gabor filter. Additionally, a change in phase may also have effects on the entire Gabor filter. The change in phase brings about a change in the waveform at the center of the convolution kernel of the Gabor filter. If the center point of the filter kernel is directly in front of the wave peak (phase 0), the filtering effect of the whole image will be enhanced. On the contrary, if the center point of the filter kernel is directly opposite to the trough (phase 180), the filtering effect will be weakened. Therefore, it is necessary to avoid coinciding the filter center point with the zero-crossing of the waveform; otherwise, the effect of the filter may not be seen. Therefore, when using the Gabor convolution kernel, it is important to adjust each parameter to the appropriate range so that the desired effect may be seen.



MRI scans from the BraTS2020 dataset were used to showcase four different imaging modalities used to examine the brain: FLAIR, T1, T1CE, and T2. Each modality provides distinct anatomical and pathological insights, which are critical for comprehensive brain tumor analysis. In brain tumor research and diagnostics, multimodal MRI is essential because each modality highlights unique tissue characteristics, enabling a more thorough assessment of tumor presence, type, and extent.

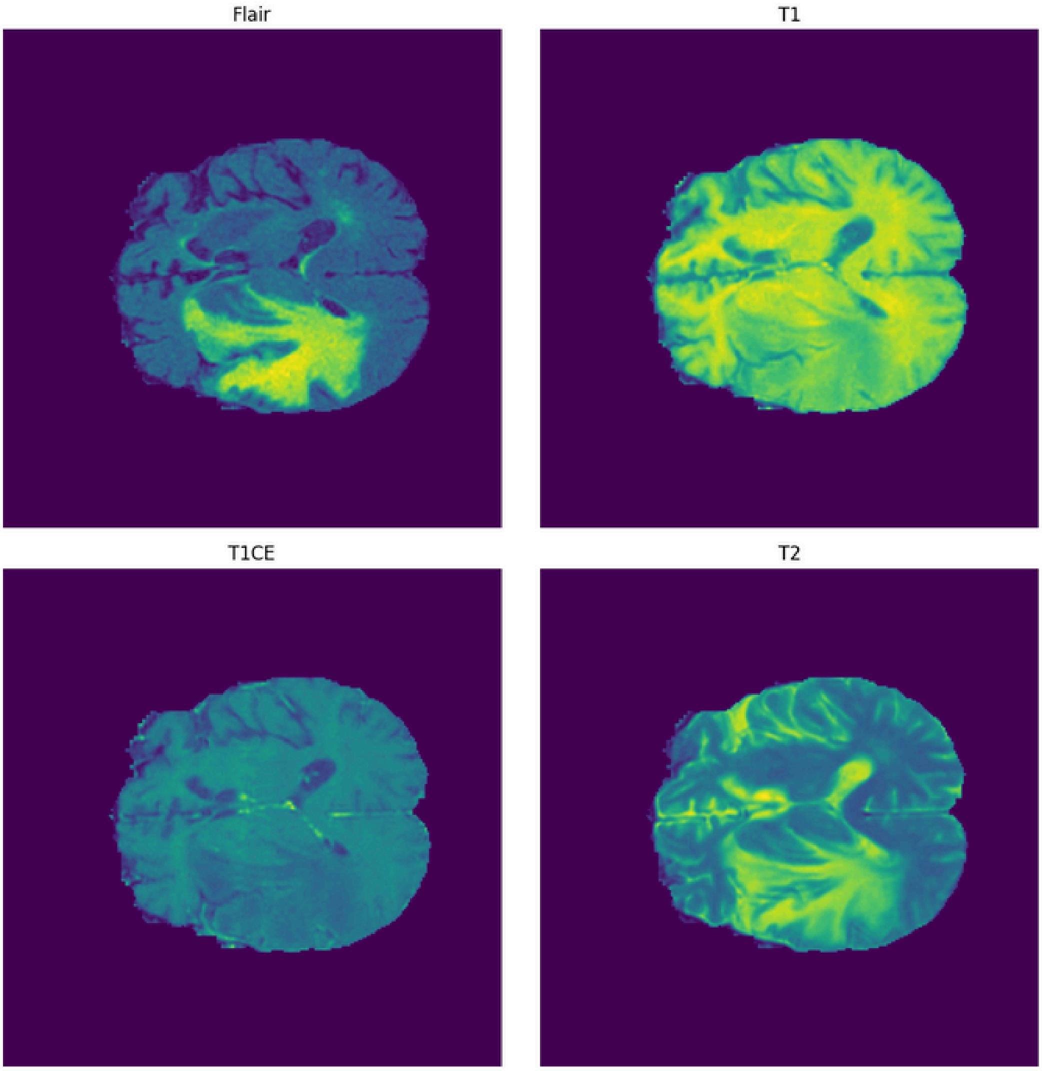
1. FLAIR (Fluid-Attenuated Inversion Recovery): FLAIR images are particularly valuable for detecting edematous regions and other abnormalities near cerebrospinal fluid (CSF) by suppressing the CSF signal, thus enhancing contrast for lesions or tissue abnormalities. In tumor analysis, FLAIR is instrumental in identifying regions with high fluid content, which often corresponds to edema surrounding the tumor.

2. T1-weighted MRI: T1 images provide high anatomical detail by capturing contrast based on the tissue's longitudinal relaxation time. T1-weighted images are effective in distinguishing normal brain structures but may not always provide optimal contrast for certain types of abnormalities. However, they serve as a baseline to observe anatomical integrity and are often paired with T1CE for enhanced analysis.

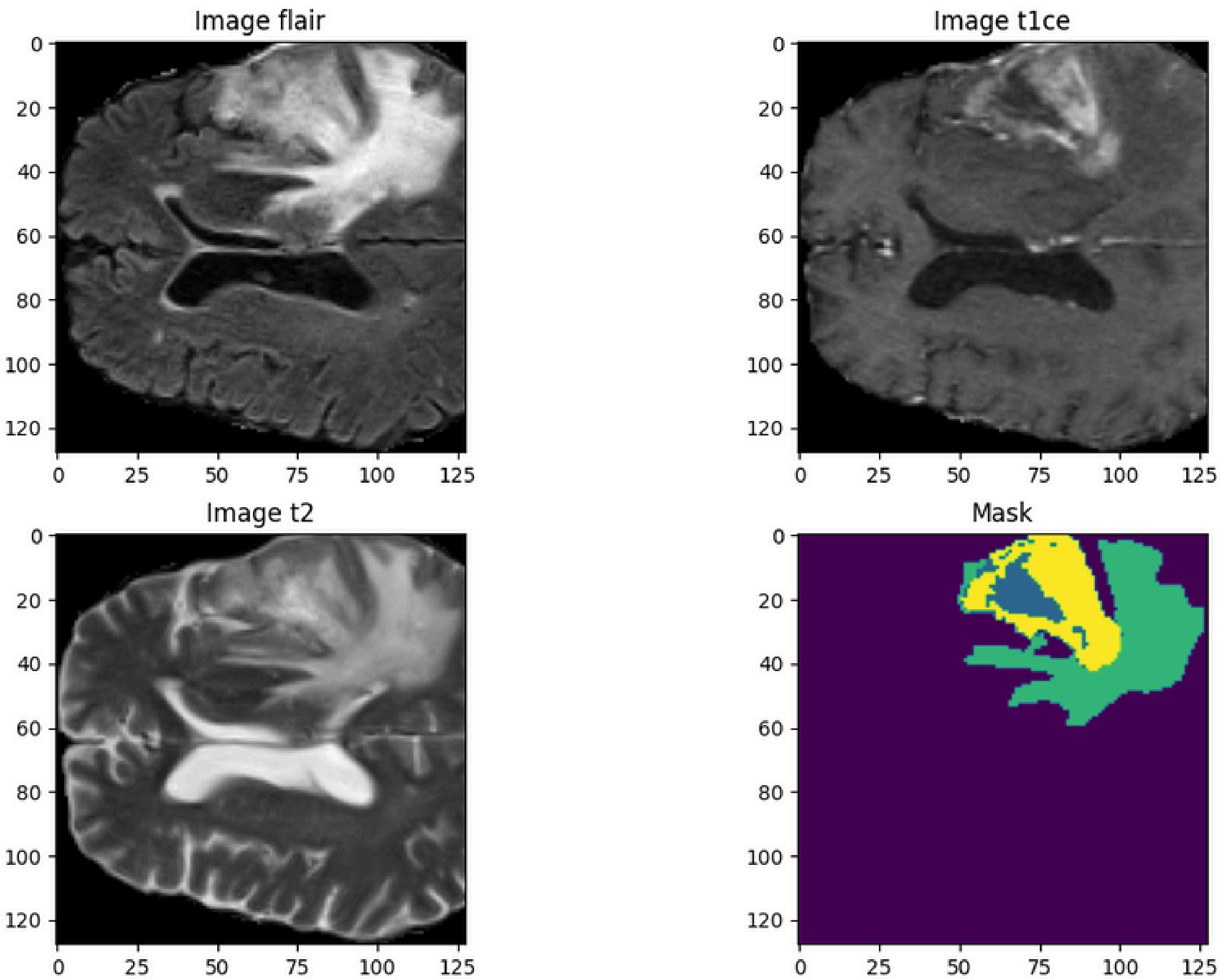
3. T1CE (T1 with Contrast Enhancement): This modality involves the injection of a gadolinium-based contrast agent, which enhances visualization of the tumor boundaries and regions with abnormal vascular permeability. In T1CE images, the contrast agent accumulates in areas with compromised blood-brain barrier integrity, such as tumor margins, making it an essential modality for delineating the tumor core and understanding the tumor’s invasive characteristics.

4. T2-weighted MRI: T2 images provide contrast based on the tissue's transverse relaxation time and are effective in detecting fluid-containing regions, making them useful for visualizing edema and differentiating between various types of tissue. T2-weighted scans are valuable for identifying brain abnormalities, as they highlight water-rich regions, which often correspond to edema or tumor infiltration.

Using these four MRI modalities in tandem provides a robust dataset for the analysis and segmentation of brain tumors. Each modality’s distinct imaging properties allow machine learning and image processing techniques, such as U-Net architectures and Gabor filters, to detect and delineate tumor regions with greater precision. In the context of brain tumor research, multimodal imaging not only facilitates accurate tumor segmentation but also enables insights into tumor morphology, growth patterns, and potential impact on surrounding brain structures. This multimodal approach, therefore, is essential for advancing both clinical and computational neuro-oncology research.



The processed image highlights the tumor region within the brain tissue, with the tumor appearing as a well-defined, hyperintense mass distinctly separated from the surrounding healthy parenchyma. By following this approach we have successfully delineated the boundaries of the tumor, enabling effective isolation of the affected area from the neighboring tissues. This precise segmentation is essential for enhancing the subsequent stages of image analysis, facilitating quantitative assessments such as volume estimation, shape characterization, and potential growth modeling. Moreover, the accuracy of this masking approach provides a robust foundation for further applications, including automated classification, tumor grading, and treatment planning, ultimately contributing to improved diagnostic precision and patient-specific therapeutic interventions.



To evaluate the performance of our U-Net model for brain tumor segmentation, we used the Intersection over Union (IoU) metric. The IoU measures the overlap between the predicted segmentation mask and the ground truth, providing a reliable indicator of segmentation accuracy. Our optimized U-Net model achieved an IoU score of 0.75, indicating a high degree of accuracy in delineating tumor boundaries. This result demonstrates the model's effectiveness in capturing the complex structures of brain tumors while minimizing false positives and false negatives.

Source Code Link:
https://colab.research.google.com/drive/1l9cogJRz897JPS2HGY37TarhM6g0qB6usp=sharing

Conclusion

Our project has demonstrated an effective multi-resolution approach for brain tumor detection by combining Gabor filters and the U-Net segmentation model, enhanced through systematic parameter optimization. The integration of Gabor filters allowed for capturing crucial texture details at multiple scales and orientations, which are essential in identifying unique patterns associated with tumor tissue. These filters were optimized to target specific spatial frequencies, making them sensitive to tumor textures and boundaries, thereby improving the accuracy of tumor region identification.

The U-Net architecture, specifically designed for segmentation tasks, proved essential in this project. Its encoder-decoder structure, with skip connections, enabled detailed feature extraction while preserving spatial information. The skip connections allowed the model to retain high-resolution details across different scales, crucial for accurate segmentation of complex structures like brain tumors. We optimized U-Net parameters, including learning rate, batch size, and filter count, to further enhance segmentation accuracy and model robustness, thereby reducing the rates of false positives and false negatives.

Training and validating this model on the BraTS2020 MRI dataset, which includes FLAIR, T1, T1CE, and T2 imaging modalities, provided a comprehensive platform for assessing the model's capabilities. Each MRI modality contributes unique insights into the tumor structure and surrounding tissue, with FLAIR identifying edematous regions, T1 providing anatomical details, T1CE emphasizing tumor boundaries, and T2 highlighting fluid-containing regions. This multimodal dataset facilitated precise tumor morphology assessment, segmentation accuracy, and a deeper understanding of tumor growth patterns and impact on neighboring brain tissue.

In conclusion, our project contributes a promising framework for automated brain tumor detection with high accuracy, sensitivity, and specificity. The combination of multi-scale Gabor filters with U-Net's segmentation ability, complemented by targeted parameter tuning, ensures a powerful and adaptable model. This approach has significant potential for applications in clinical diagnostics, aiding radiologists in tumor detection, and enabling patient-specific treatment planning. The model's generalizability also makes it well-suited for broader applications in medical image analysis, where multi-scale feature detection and segmentation are vital for early diagnosis and therapeutic intervention.

References:

- Rabah Hammouche, Abdelouahab Attia, Samir Akhrouf, Zahid Akhtar, Gabor filter bank with deep autoencoder based face recognition system, Expert Systems with Applications, Volume 197, 2022, 116743, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2022.116743. (https://www.sciencedirect.com/science/article/pii/S0957417422002123)

- V. Damerjian, O. Tankyevych, N. Souag, E. Petit, Speckle characterization methods in ultrasound images – A review, IRBM, Volume 35, Issue 4, 2014, Pages 202-213, ISSN 1959-0318, https://doi.org/10.1016/j.irbm.2014.05.003. (https://www.sciencedirect.com/science/article/pii/S1959031814000797)

- Qian Z, Metaxas DN, Axel L. Extraction and tracking of MRI tagging sheets using a 3D Gabor filter bank. Conf Proc IEEE Eng Med Biol Soc. 2006;2006:711-4. doi: 10.1109/IEMBS.2006.259542. PMID: 17946853.

- Cao, Yuan, and Yinglei Song. 2024. "New Approach for Brain Tumor Segmentation Based on Gabor Convolution and Attention Mechanism" Applied Sciences 14, no. 11: 4919. https://doi.org/10.3390/app14114919

- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234–241.https://arxiv.org/abs/1505.04597

- Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., & Jorge Cardoso, M. (2017). Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 240–248

- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & van Ginneken, B. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88.

- Bhattacharya, Debmalya & Jibanpriya, Mrs & Bhattacherjee, Payal. (2013). Brain Image Segmentation Technique Using Gabor filter parameter.